



# Skew Generalized Extreme Value Distribution: Probability Weighted Moments Estimation and Application to Block Maxima Procedure

Pierre Ribereau, Esterina Masiello, Philippe Naveau

## ► To cite this version:

Pierre Ribereau, Esterina Masiello, Philippe Naveau. Skew Generalized Extreme Value Distribution: Probability Weighted Moments Estimation and Application to Block Maxima Procedure. Communications in Statistics - Theory and Methods, 2014, pp.1-25. hal-01018877

**HAL Id: hal-01018877**

**<https://hal.science/hal-01018877>**

Submitted on 7 Jul 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Skew Generalized Extreme Value Distribution: Probability Weighted Moments Estimation and Application to Block Maxima Procedure

Pierre Ribereau<sup>(\*)</sup>, Esterina Masiello<sup>(\*)</sup>, Philippe Naveau<sup>(\*\*)</sup>

<sup>(\*)</sup> Université de Lyon  
Université Lyon 1  
CNRS, UMR5208, Institut Camille Jordan  
43 blvd du 11 novembre 1918  
F-69622 Villeurbanne-Cedex, France

<sup>(\*\*)</sup> Laboratoire des Sciences du Climat et de l'Environnement,  
IPSL-CNRS  
Orme des Merisiers / Bat. 709 C.E. Saclay  
91191 Gif-sur-Yvette, FRANCE

**Abstract.** Following the work of Azzalini ([2] and [3]) on the skew normal distribution, we propose an extension of the Generalized Extreme Value (GEV) distribution, the SGEV. This new distribution allows for a better fit of maxima and can be interpreted as both the distribution of maxima when maxima are taken on dependent data and when maxima are taken over a random block size. We propose to estimate the parameters of the SGEV distribution via the Probability Weighted Moments method. A simulation study is presented to provide an application of the SGEV on block maxima procedure and return level estimation. The proposed method is also implemented on a real-life data.

**AMS Subject classifications:** 62F10, 62E17, 62P12.

**Key words and phrases:** Skew distributions, Generalized Extreme Value distribution, Extreme Value Theory, Probability Weighted Moments, Return Level Estimation.

# 1 Skew distributions and Extreme Value Theory

The idea of modifying a symmetric probability density function of a random variable (r.v.) in a multiplicative fashion, introducing skewness, has been implemented for the univariate normal distribution by Azzalini ([2], [3]) yielding the so-called skew-normal distribution having the probability density function of the form  $2\phi(x)\Phi(\lambda x)$ ,  $x, \lambda \in \mathbb{R}$  (where  $\phi(\cdot)$  and  $\Phi(\cdot)$  denote respectively the probability distribution function (pdf) and cumulative distribution function (cdf) of the standard normal distribution). This distribution presents the advantage to look like a normal one in the core data but allowing some asymmetry. The model has since been extended to the gaussian related distributions (multivariate normal or Student distribution) and its extensions are provided in the book edited by Genton [12] and in the review paper [4] and the monograph by Azzalini [5] who regularly updates list of references on his web page. The skew-normal distribution has been also applied in many fields like climatology (Flecher *et al.* [11]) or insurance (Vernic [28]) and adapted to classical statistical tools like Space-state models (Naveau *et al.* [24]) or random fields (Allard and Naveau [1]). Since then, several authors have tried to generalize these results to skewing arbitrary symmetric (or not) pdf's. See Gupta et al. [15] for a most detailed discussion of skew-symmetric models based on the normal, Student's t, Cauchy, Laplace, logistic and uniform distributions. In that case, the main feature of these models is that a new parameter  $\lambda$  is introduced to control skewness and kurtosis. Nadarajah and Kotz ([21], [22] and [23]) proposed some skew models based on Cauchy, Laplace, Logistic, Bessel and Pearson (among others).

Some relations between Extreme Value Theory and Skew Distribution have already been studied by Chang and Genton [7] in the case of skew-symmetric family of distributions and by Lysenko *et al.* [20] in the case of multivariate generalized skew-normal distribution.

In the same spirit, we propose here an extension of the Generalized Extreme Value (GEV) distribution. This distribution, introduced by Jenkinson [19], is the limiting distribution for extremes, both normalized maxima and minima of a random sample. In a more formal way, it is known that the only possible limiting form of a normalized maximum of a random sample

when a non-degenerate limit exists is captured by the GEV

$$G(x; \sigma, \gamma, \mu) = \begin{cases} \exp\left[-(1 + \gamma \frac{x-\mu}{\sigma})^{-1/\gamma}\right], & \text{if } 1 + \gamma \frac{x-\mu}{\sigma} > 0, \gamma \neq 0, \\ \exp\left[-\exp(-\frac{x-\mu}{\sigma})\right], & \text{if } x \in \mathbb{R}, \gamma = 0, \end{cases}$$

where  $\mu \in \mathbb{R}$ ,  $\sigma > 0$  and  $\gamma$  are called the location, scale and shape parameters, respectively.

The GEV model has been applied to environmental sciences to fit, for example, yearly maximal temperatures or yearly maximal river discharges. However, the limiting distribution does not always provide a good fit in the finite samples that arise in practice, since the extreme value theory is based on asymptotic arguments.

Our purpose, here, is to generalize the GEV distribution by addition of an extra parameter,  $\lambda$ , introduced to model flexibility in the data and so better adapt to real situations. The main features of this distribution will be discussed together with two methods of estimation of its parameters.

The paper is organized as follows. In Section 2, we will define the distribution, study some of its characteristics and present some possible interpretations. In Section 3, the parameters of the distribution will be estimated by the classical methods of maximum likelihood and probability-weighted moments. Finally, a simulation study and an application to real data will be performed (Section 4 and 5) to investigate the behavior of the estimator in small and moderate samples.

## 2 The skew generalized extreme value distribution

**Definition 1.** *A random variable  $X$  is said to have a Skew Generalized Extreme Value (SGEV) distribution if its probability density function is given by*

$$f(x) = (\lambda + 1)g(x)G^\lambda(x) \quad (1)$$

where  $\lambda > -1$ ,  $g(\cdot)$  and  $G(\cdot)$  are respectively the density function and the cumulative distribution function of a GEV distribution with location parameter  $\mu$ , scale parameter  $\sigma$  and shape parameter  $\gamma$ .

The cumulative distribution function is given by

$$F(x) = G^{\lambda+1}(x).$$

The distribution will be denoted for short by  $SGEV(\mu, \sigma, \gamma, \lambda)$ . It follows from (1) that the pdf of  $X$  can be written as

$$f(x) = \begin{cases} \left(\frac{\lambda+1}{\sigma}\right) \left(1 + \gamma \frac{x-\mu}{\sigma}\right)^{-1/\gamma-1} \exp \left[ -(\lambda+1) \left(1 + \gamma \frac{x-\mu}{\sigma}\right)^{-1/\gamma} \right] & \text{if } 1 + \gamma \frac{x-\mu}{\sigma} > 0, \gamma \neq 0, \\ \left(\frac{\lambda+1}{\sigma}\right) \exp \left[ -(\lambda+1) \exp \left( -\frac{x-\mu}{\sigma} \right) \right] \exp \left( -\frac{x-\mu}{\sigma} \right) & \text{if } x \in \mathbb{R}, \gamma = 0. \end{cases}$$

When  $\lambda = 0$ , equation (1) reduces to the GEV pdf. The parameter  $\lambda$  allows for a greater degree of flexibility and we can expect this to be useful in many more practical situations.

The quantile function of the cdf  $F$  is given by

$$Q(u) = \begin{cases} \mu + \frac{\sigma}{\gamma} \left( \left( \frac{-\log u}{\lambda+1} \right)^{-\gamma} - 1 \right) & \text{if } \gamma \neq 0, \\ \mu - \sigma \log \left( \frac{-\log u}{\lambda+1} \right) & \text{if } \gamma = 0 \end{cases}$$

and allows to compute quantiles in an easy way.

We also find that the expectation, the second-order moment and the variance of  $X$  are respectively given by

$$\mathbb{E}(X) = \mu - \frac{\sigma}{\gamma} [1 - (\lambda+1)^\gamma \Gamma(1-\gamma)],$$

$$E(X^2) = \left[ \left( \mu - \frac{\sigma}{\gamma} \right) + \left( \frac{\sigma}{\gamma(\lambda+1)^{-\gamma}} \Gamma(1-\gamma) \right) \right]^2,$$

$$V(X) = \frac{\sigma^2}{\gamma^2(\lambda+1)^{-2\gamma}} [\Gamma(1-2\gamma) - \Gamma^2(1-\gamma)],$$

if  $\lambda > -1$ , where  $\Gamma$  denotes the Gamma function  $\Gamma(t) = \int_0^\infty e^{-u} u^{t-1} du$ ,  $t > 0$ . Simple modifications of Chang and Genton [7] leads to the conclusion that the extreme value index of the SGEV distribution is  $\gamma$ .

The SGEV distribution offers some interesting practical interpretations which are going to be presented in the following remarks.

**Remark 1.** Let  $X_1, X_2, \dots$  be i.i.d. r.v.'s with cdf  $F$  and let  $N$  be a Poisson r.v. with intensity  $\lambda^* > 0$ . Moreover, let  $N$  be independent of  $\{X_i\}_{i \geq 1}$ . Define the following maximum with a random block size

$$M = \begin{cases} \max(X_1, \dots, X_N) & \text{if } N \geq 1, \\ 0 & \text{if } N = 0. \end{cases}$$

Then, we have

$$\begin{aligned} P(M \leq x) &= \sum_{n=0}^{+\infty} F^n(x) \frac{\lambda^{*n} e^{-\lambda^*}}{n!} \\ &= \exp\{-\lambda^* [1 - F(x)]\} \\ &= \{\exp[-\bar{F}(x)]\}^{\lambda^*}. \end{aligned}$$

Now, denote  $G(x) = \exp[-\bar{F}(x)]$ . It follows

$$P(M \leq x) = G^{\lambda^*}(x) \quad (2)$$

which corresponds to SGEVs with strictly positive parameters  $\lambda^* = \lambda + 1$ .

Define the cdf  $H(x; \sigma, \gamma, \mu)$  by

$$H(x; \sigma, \gamma, \mu) = \begin{cases} 1 - \{1 + \gamma(\frac{x-\mu}{\sigma})\}^{-1/\gamma}, & \text{if } \gamma \neq 0, \\ 1 - \exp[-(\frac{x-\mu}{\sigma})], & \text{if } \gamma = 0, \end{cases}$$

where  $x \geq \mu$  if  $\gamma \geq 0$ ,  $\mu \leq x \leq \mu - \sigma/\gamma$  if  $\gamma < 0$  and  $\mu \in \mathbb{R}$  is the location parameter,  $\sigma > 0$  the scale parameter and  $\gamma \in \mathbb{R}$  the shape parameter. This distribution, which appears as the limit distribution of scaled excesses over high thresholds, is called a Generalized Pareto Distribution (GPD) (Pickands [25]).

If  $\bar{F}$  is a GPD tail, then the cdf of  $M$  is the cdf of a SGEV on the interval  $[\mu, \infty[$ , as it is shown in Figure 1. If  $x \leq \mu$ , then  $\mathbb{P}(M \leq x) = \mathbb{P}(N = 0) = \exp(-\lambda^*)$ . The equality (2) offers then a first interpretation for the SGEV distribution: the cdf  $G^{\lambda^*}(x)$ ,  $\lambda^* = \lambda + 1$ , on the interval  $[\mu, \infty[$  can be viewed as the cdf of a maximum taken over exceedances in a random Poisson block size.

**Remark 2.** The SGEV is also the distribution of the maxima when they are taken on dependent data. In that case,  $\lambda$  is between  $-1$  and  $0$ . The extremal index  $\theta$  (see, for example, Beirlant et al. [6]) is then equal to

$\lambda + 1$ . The SGEV is more general since the value of  $\lambda$  needs only to be greater than  $-1$ . So, the family of SGEV includes the classical GEV one, but also the GEV one based on dependent data. This is quite interesting because this property could be used to define a new plug-in estimator of the extremal index,  $\hat{\theta} = \hat{\lambda} + 1$ , where  $\hat{\lambda}$  is the estimator of  $\lambda$  obtained by the PWM method.

### 3 Parameter estimation

Let  $(X_1, \dots, X_n)$  be a sample of SGEV r.v.s. In this section, the parameters  $(\mu, \sigma, \gamma, \lambda)$  of the SGEV distribution will be estimated by both the method of maximum likelihood (ML) and the probability weighted moments (PWM) method.

#### 3.1 Maximum likelihood

In case  $\gamma \neq 0$ , the log-likelihood function for a sample  $(X_1, \dots, X_n)$  of SGEV r.v.s is given by

$$\begin{aligned} \log L(\mu, \sigma, \gamma, \lambda) &= n \log(\lambda + 1) - n \log \sigma \\ &\quad - \left( \frac{1}{\gamma} + 1 \right) \sum_{i=1}^n \log \left( 1 + \gamma \frac{x_i - \mu}{\sigma} \right) \\ &\quad - (\lambda + 1) \sum_{i=1}^n \left( 1 + \gamma \frac{x_i - \mu}{\sigma} \right)^{-1/\gamma}. \end{aligned}$$

If  $\gamma = 0$ , the log-likelihood function reduces to

$$\begin{aligned} \log L(\mu, \sigma, \lambda) &= n \log(\lambda + 1) - n \log \sigma - (\lambda + 1) \sum_{i=1}^n \exp \left( -\frac{x_i - \mu}{\sigma} \right) \\ &\quad - \sum_{i=1}^n \left( \frac{x_i - \mu}{\sigma} \right). \end{aligned}$$

The ML estimator  $(\hat{\mu}, \hat{\sigma}, \hat{\gamma}, \hat{\lambda})$  is obtained by maximizing the log-likelihood function. Clearly, no explicit solution exists to these equations so that numerical procedures are called for. For a discussion about the properties of

the MLE see Smith [27]. In the paper by Smith [27], the ML estimators of the parameters of the GEV distribution are shown to be asymptotically normal whenever  $\gamma > -1/2$ . Since the SGEV distribution is based on the GEV distribution, it is easy to check that the classical properties of the ML estimator of  $(\mu, \sigma, \gamma, \lambda)$  hold whenever  $\gamma > -1/2$ .

**Remark 3.** *The ML method usually gives poor results when it has to estimate a skew parameter: for example, for moderate sample sizes, the ML estimator is infinite with positive probability when dealing with skew normal distribution (Sartori [26]). The issue mentioned here has a close connection with the Fisher singularity problem discussed in Hallin and Ley (2012a, 2012b).*

*In Section 4.1 and in Tables 1 to 4, the results obtained by estimating the parameters by Maximum Likelihood are disappointing. This is the reason why, when estimating the return level in Section 4.2, we will avoid the ML method and prefer the Probability Weighted Moments one, which will be described in the next section.*

### 3.2 The method of probability weighted moments

The probability weighted moments, introduced by Greenwood *et al.* [13], are a simple generalization of the usual moments of a probability distribution. They are defined in terms of the cdf  $H$  of a r.v.  $Y$  as

$$\mathbb{E}\{Y^p[H(Y)]^r[1 - H(Y)]^s\}$$

with  $p$ ,  $r$  and  $s$  real numbers. Particularly useful special cases are the probability-weighted moments

$$\beta_r = \mathbb{E}\{Y[H(Y)]^r\}$$

which have been used by Hosking *et al.* (1985) to estimate the parameters of the GEV distribution. For the GEV distribution, they equal

$$\beta_r = (r + 1)^{-1} \left\{ \mu - \frac{\sigma}{\gamma} \left[ 1 - \frac{\Gamma(1 - \gamma)}{(1 + r)^{-\gamma}} \right] \right\} \quad \gamma < 1, \gamma \neq 0. \quad (3)$$

This method of moments has been also applied to the truncated skew-normal distribution (Flecher *et al.* [10]). It can be equally applied to estimate the parameters of the SGEV distribution. Let  $X$  be a  $SGEV(\mu, \sigma, \gamma, \lambda)$  r.v. with cdf  $F$  and let  $Y$  be a  $GEV(\mu, \sigma, \gamma)$  r.v. with cdf  $G$ . It is easy



to check that the probability weighted moments of a SGEV distribution, denoted by  $\nu_a$ , are special probability weighted moments of a GEV distribution:

$$\begin{aligned}
\nu_a &= \mathbb{E}[XF^a(X)] \\
&= \int xF^a(x)f(x)dx \\
&= \int xG^{(\lambda+1)a}(x)(\lambda+1)g(x)G^\lambda(x)dx \\
&= (\lambda+1) \int xG^{(\lambda+1)a+\lambda}(x)g(x)dx \\
&= (\lambda+1)\mathbb{E}[YG^{\lambda(a+1)+a}(Y)] = (\lambda+1)\beta_{\lambda(a+1)+a}
\end{aligned}$$

By taking into account (3), we obtain

$$\nu_a = \mathbb{E}[XF^a(X)] = \frac{1}{a+1} \left\{ \mu - \frac{\sigma}{\gamma} \left[ 1 - \frac{\Gamma(1-\gamma)}{((\lambda+1)(a+1))^{-\gamma}} \right] \right\}.$$

In practice, we just need to choose four different values of  $a$  and so set up the necessary number of moment equations. Here, we choose  $a = 0, 1, 2, 3$ , because they are trivial to estimate. The estimators of  $\gamma$ ,  $\sigma$ ,  $\mu$  and  $\lambda$  are then obtained by solving the following system of equations:

$$\left\{ \begin{array}{lcl} \nu_0 & = & \mu - \frac{\sigma}{\gamma} [1 - (\lambda+1)^\gamma \Gamma(1-\gamma)], \\ 2\nu_1 & = & \left\{ \mu - \frac{\sigma}{\gamma} [1 - 2^\gamma (\lambda+1)^\gamma \Gamma(1-\gamma)] \right\}, \\ 3\nu_2 & = & \left\{ \mu - \frac{\sigma}{\gamma} [1 - 3^\gamma (\lambda+1)^\gamma \Gamma(1-\gamma)] \right\}, \\ 4\nu_3 & = & \left\{ \mu - \frac{\sigma}{\gamma} [1 - 4^\gamma (\lambda+1)^\gamma \Gamma(1-\gamma)] \right\}, \end{array} \right. \quad (4)$$

after replacing  $\nu_a$  by its empirical version  $\hat{\nu}_a$ :

$$\hat{\nu}_a = \frac{1}{n} \sum_{i=1}^n X_{(i)} \left( \frac{i}{n} \right)^a$$

where  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  represent the ordered sample. This system is quiet hard to solve numerically.

A promising avenue consists of first estimating  $\gamma$  by solving the following equation

$$\frac{3\hat{\nu}_2 - \hat{\nu}_0}{2\hat{\nu}_1 - \hat{\nu}_0} = \frac{3^{\hat{\gamma}} - 1}{2^{\hat{\gamma}} - 1}. \quad (5)$$

Once  $\hat{\gamma}$  given, find the estimations of  $\lambda$  and  $\sigma$  by solving the following system of equations

$$\begin{cases} 2\hat{\nu}_1 - \hat{\nu}_0 &= \frac{\hat{\sigma}}{\hat{\gamma}} \left\{ \Gamma(1 - \hat{\gamma})(\hat{\lambda} + 1)^{\hat{\gamma}}(2^{\hat{\gamma}} - 1) \right\}, \\ 4\hat{\nu}_3 - 3\hat{\nu}_2 &= \frac{\hat{\sigma}}{\hat{\gamma}} \left\{ \Gamma(1 - \hat{\gamma})(\hat{\lambda} + 1)^{\hat{\gamma}}(4^{\hat{\gamma}} - 3^{\hat{\gamma}}) \right\}, \end{cases} \quad (6)$$

and finally,  $\hat{\sigma}$ ,  $\hat{\gamma}$  and  $\hat{\lambda}$  given, obtain the estimation of  $\mu$ , that is

$$\hat{\mu} = \hat{\nu}_0 + \frac{\hat{\sigma}}{\hat{\gamma}} \left[ 1 - (\hat{\lambda} + 1)^{\hat{\gamma}} \Gamma(1 - \hat{\gamma}) \right].$$

The case  $\gamma = 0$  can also be covered by this method: a system of three equations with three unknown parameters will then be obtained by taking the first three equations in the system (4) and letting  $\gamma$  go to 0. Nevertheless, in the case  $\gamma = 0$ , all moments exist and then, we do not need a new method based on moments to get good results since the classical moments method will be satisfactory.

To derive the limiting distribution of  $(\hat{\mu}, \hat{\sigma}, \hat{\gamma}, \hat{\lambda})$ , we need to study the limiting behavior of  $\hat{\nu}_a$ . By direct application of the central limit theorem, one can easily prove that  $\hat{\nu}_a$  is asymptotically normally distributed with mean  $\nu_a$  and asymptotic variance

$$\frac{(\lambda + 1)^2}{n} \left\{ \frac{\sigma^2}{\gamma^2} [(a + 1)(\lambda + 1)]^{2\gamma} \left[ \Gamma(1 - 2\gamma) K \left( \frac{a(\lambda + 1) + \lambda}{(a + 1)(\lambda + 1)} \right) - \Gamma^2(1 - \gamma) \right] \right\}$$

where  $K$  denotes the hypergeometric function  $K(x) = {}_1F_2(k, 2k; 1 + k; -x)$ . A similar argument applies to any linear combination of the  $\hat{\nu}_a$  ( $a = 0, 1, 2, \dots$ ) and it follows that the  $\hat{\nu}_a$  are asymptotically jointly normal.

**Theorem 1.** *Let  $(X_1, \dots, X_n)$  be a sample of SGEV. If  $\gamma < 0.5$ , then the rescaled PWM estimator vector defined by (4) and denoted by*

$$\sqrt{n} \begin{pmatrix} \hat{\nu}_0 \\ \hat{\nu}_1 \\ \hat{\nu}_2 \\ \hat{\nu}_3 \end{pmatrix}$$

converges in distribution, as  $n \rightarrow \infty$ , towards

$$\mathcal{N} \left( \begin{pmatrix} \nu_0 \\ \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \Omega \right)$$

where  $\Omega$  denotes the covariance matrix.

The elements of the covariance matrix  $\Omega$  are given in Appendix A.

**Corollary 2.** *The estimators of  $\mu$ ,  $\sigma$ ,  $\gamma$  and  $\lambda$  obtained by the probability weighted moments method are asymptotically normally distributed and are asymptotically unbiased if  $\gamma < 0.5$ .*

The proof of the corollary is a direct application of Theorem 1 and the Delta Method.

**Remark 4.** *Moments  $\nu_0, \dots, \nu_3$  have been chosen in the estimation procedure of the parameters of the SGEV distribution. The same moments could be used to obtain different moment methods. Equations similar to (5) could be written down with any of the 4 sets of 3 moments, among  $\nu_0, \dots, \nu_3$ . Similarly, any pair of moments among the 6 possible pairs can lead to a system of two equations similar to that in (6). Finally,  $\mu$  can also be estimated using any of the 4 moments. The estimation variance of the selected combination of moments is the lowest among all possible combinations, since the moments of lowest degree have been selected. This procedure can then be improved using all combinations simultaneously. Consider for example the estimation of  $\gamma$ . Four equations similar to (5) can be written using each triplet of moments. An estimator  $\hat{\gamma}$  can be defined as the value minimizing the mean squared errors between the left hand side and the right hand side of these equations. A similar approach can then be implemented for estimating  $\sigma$  and  $\lambda$  given  $\hat{\gamma}$ , and finally for estimating  $\mu$  given all other estimates.*

## 4 A simulation study

### 4.1 Parameters estimation

We perform a simulation study in order to compare the performances of the ML and PWM estimators of the parameters of the SGEV distribution.

Simulations were performed for sample size  $n = 100$  with  $\gamma$  taking values  $-0.5, 0, 0.5, 1$  and  $1.5$ . Without loss of generality, the location and scale parameters were set to  $\mu = 0$  and  $\sigma = 1$ . We set also the skew parameter  $\lambda$  to  $0.2, 0.5, 1$  and  $1.5$ . For each combination of values of  $\gamma$  and  $\lambda$ , we generated 100 random samples from the SGEV distribution and for each sample, we estimate the parameters  $\mu, \sigma, \gamma$  and  $\lambda$  by ML and PWM method. The results are listed in Tables 1 to 4.

Remark that some chosen values for  $\gamma$  are not in the range of validity of the estimation methods. In practice, on real data, the value of  $\gamma$  is unknown, so values which are not in the validity range were used to discuss the robustness of the method. For example, when  $\gamma$  is greater than 0.5 but lesser than 1, moments still exist but the estimators are no more asymptotically normal. In this case, the PWM estimator gives acceptable results when estimating  $\mu$  and  $\sigma$  (sometimes  $\lambda$ ). On the other side, when  $\gamma \geq 1$  the PWM method is not satisfactory anymore.

As previously mentioned, the Maximum Likelihood method does not perform well especially when estimating the skew parameter  $\lambda$ . Indeed, in the case  $\gamma < 0.5$ , we can notice the superiority of the PWM method for estimating the skew parameter  $\lambda$ , while ML and PWM provides good and comparable estimations for the three other parameters. In particular, because of the quite large sample size ( $n = 100$ ) in our simulation study, the ML estimator is preferable to estimate  $\gamma$  (when  $\gamma < 0.5$ ). On the other side, both methods do not perform well when  $\gamma$  gets greater (if  $\gamma > 0.5$ ), especially for the skew parameter. For the PWM method, this is due to the usual regularity conditions underlying the asymptotic properties of the estimators ( $\gamma < 0.5$ ) while for the ML method this may be due to some algorithmic problems. This shows that both methods are not really robust when used outside their range of validity. For a more detailed discussion about the performance of both ML and PWM estimators, see Hosking and Wallis [18].

**Remark 5.** *When looking at the results of the simulation study, it appears that the parameter  $\gamma$  is always fairly well estimated by ML. This could suggest to consider a mixed estimation procedure which will consist in estimating  $\gamma$  with the ML method and the other parameters with the PWM method. This kind of procedure would lead to a vector of estimators whose asymptotic law could be quite hard to obtain. An alternative could be to consider the estimation obtained by the PWM method as a starting value for the ML method.*

## 4.2 Return level estimation

Let  $Z_1, \dots, Z_n$  be a sample of i.i.d. random variables from a distribution function  $H$ . We are now interested in estimating  $z_p := H^{-1}(1 - p)$ . The method of Block Maxima consists in replacing the original sample  $Z_1, \dots, Z_n$  by  $Y_1, \dots, Y_m$  where  $Y_i$  is the maximum on the  $i^{th}$  block of size  $k$  i.e.  $Y_i = \max\{Z_{(i-1)k+1}, \dots, Z_{ik}\}$  with  $n = m \times k$ . Since the r.v.'s  $Y_i$  are asymptotically GEV distributed, the return level can be estimated using (see Beirlant *et al.* [6])

$$\hat{z}_p^* = \begin{cases} \mu_Y + \frac{\sigma_Y}{\gamma_Y} \left[ (-\log(1 - p)^k)^{-\gamma_Y} - 1 \right], & \text{if } \gamma_Y \neq 0, \\ \mu_Y - \sigma_Y \log(-\log(1 - p)^k), & \text{if } \gamma_Y = 0, \end{cases} \quad (7)$$

leading to the GEV estimator  $\hat{z}_p^{GEV}$  after replacing  $\mu_Y$ ,  $\sigma_Y$  and  $\gamma_Y$  by  $\hat{\mu}_Y$ ,  $\hat{\sigma}_Y$  and  $\hat{\gamma}_Y$ , the ML or PWM estimators of the GEV parameters based on the sample  $Y_1, \dots, Y_m$ .

If  $Z_1, \dots, Z_n$  are i.i.d. with gaussian distribution  $\mathcal{N}(0, 1)$  we also have (see Embrechts *et al.* [9])

$$\mathbb{P} \left( \frac{Z_{(n,n)} - d_n^N}{c_n^N} \leq z \right) \rightarrow GEV(\mu = 0, \sigma = 1, \gamma = 0)$$

where  $Z_{(n,n)}$  is the maximum of the sample,

$$c_n^N = (2 \log n)^{-1/2},$$

$$\text{and } d_n^N = \sqrt{2 \log n} - \frac{\log(4\pi) + \log \log n}{2(2 \log n)^{1/2}}.$$

The distribution of  $Y_i$  can be approximated by a  $GEV(\mu = d_k^N, \sigma = c_k^N, \gamma = 0)$  and the return level of probability  $p$  can be approximated by  $z_p^{BM}$  after replacing  $\mu_Y$  and  $\sigma_Y$  by  $d_k^N$  and  $c_k^N$ , and  $\gamma_Y$  by 0 in Equation (7) (in the case  $\gamma_Y = 0$ ).

Obviously, the theoretical estimation  $z_p^{BM}$  will get closer to  $z_p$  as the block size increases. In practice, we need to choose a block size large enough to ensure a good approximation of the GEV distribution but allowing also

enough maxima observations to obtain a good parameter estimation. In that way, the flexibility of the SGEV distribution can give a good compromise. Once the sample of maxima obtained, the return level can be estimated by

$$\hat{z}_p^{SGEV} = \begin{cases} \mu_s + \frac{\sigma_s}{\gamma_s} \left[ (-\log(1-p)^{k/(\lambda_s+1)})^{-\gamma_s} - 1 \right], & \text{if } \gamma_s \neq 0, \\ \mu_s - \sigma_s \log \left( -\log(1-p)^{k/(\lambda_s+1)} \right), & \text{if } \gamma_s = 0, \end{cases} \quad (8)$$

where  $\mu_s$ ,  $\sigma_s$ ,  $\gamma_s$  et  $\lambda_s$  are the parameter estimators obtained after fitting a SGEV on the sample of maxima  $Y_1, \dots, Y_m$ .

Similar results stand for the Log-gamma distribution  $L\Gamma(\alpha, \beta)$  with

$$c_n^{L\Gamma} = \left( (\Gamma(\beta))^{-1} (\log n)^{\beta-1} n \right)^{1/\alpha} \text{ and } d_n = 0.$$

We want to answer the two following questions : in the case of slow convergence of the maximum towards the GEV distribution (e.g. the Gaussian case), how the SGEV behaves and how does the SGEV for heavy tailed distributions?

Figures 2 and 3 represent the estimation of the quantile as a function of the block size. For each chosen distribution ( $\mathcal{N}(0, 1)$  in Figure 2 and Log-gamma(2,2) in Figure 3), we simulated 500 samples of size  $n = 500$  or  $n = 5000$ . For each combination, we compute the theoretical value  $z_p^{BM}$  (red line) to show how good the approximation by a GEV distribution is, the estimation of the quantile based on the maximum by block using both GEV ( $\hat{z}_p^{GEV}$ , grey line) and SGEV ( $\hat{z}_p^{SGEV}$ , black line) and the 95% empirical confidence intervals (dashed lines). The true value of the quantile is represented by the horizontal line. We also represent, on the bottom row of each figure, the value of the skew parameter estimation  $\lambda_s$ , together with a 95% empirical confidence interval.

By a first look at these simulations, we observe that the return level estimation is not so good when dealing with Gaussian distributions (cf Figure 2). This is not surprising since it is well known that the normal distribution converges very slowly to the GEV distribution, but it is also an indicator of the limitations of the Extreme Value Theory. Although the SGEV approximation is not great, it seems to propose an improvement in terms of bias

and variance.

When dealing with a heavy tailed distribution (cf Figure 3), the quality of the return level estimation using the SGEV approximation gets better even if the approximation of the sample of maxima seems good (see the theoretical  $z_p^{BM}$ ).

In view of these simulations, the SGEV proposes a complementary method to the classical Extreme Value Theory by improving the return level estimation both in the case of a slow convergence or in the heavy tailed case.

## 5 An application to real data

We apply the proposed procedure to the Fort Collins precipitation data. The data consist of the annual daily precipitation maxima (in mm) recorded by a weather station from 1948 to 2001 in Fort Collins (Colorado, USA) (cf Guillou *et al.* [14] or Diebolt *et al.* [8] for more details and explanation on this data set).

The values of the estimated parameters, obtained by ML and PWM methods, are listed in Table 5 when fitting a SGEV distribution or a GEV distribution. For the GEV distribution, the estimation of the return level associated to a period of 100 years is equal to 573.72 (when adopting the PWM method) and 653.33 (when adopting the ML method). By fitting the SGEV distribution, we obtained an estimation of the return level period equal to 608.83 (with the PWM method) and 653.21 (with the ML method).

This application was selected to illustrate a number of practical points. First of all, we remark that if we look at the same estimation method (either ML or PWM method) the return level estimation does not vary so much if we take a GEV or a SGEV tail approximation. This is probably due to the fact that the annual maximum is taken on a sample of important size (equal to 365) and so the GEV approximation is good enough to provide interesting results. Indeed, the values of  $\lambda$  are near 0 indicating that the GEV distribution may be a good choice of fit for this sample.

In addition, we remark that there is no loss using SGEV instead of GEV distribution. So, in our opinion, the first one should be preferred since it englobes more cases than the classical GEV distribution. Indeed, the SGEV distribution could be used to define the right block size. One can try different block size and choose the one which gives the value of the parameter

Parameter	True value	Mean for ML	Mean for PWM	sd for ML	sd for PWM
$\mu$	0	0.184	-0.014	0.452	0.103
$\sigma$	1	0.908	0.953	0.288	0.080
$\gamma$	-0.5	-0.517	-0.433	0.065	0.075
$\lambda$	0.2	0.038	0.213	0.620	0.028
$\mu$	0	0.097	0.009	0.075	0.105
$\sigma$	1	1.020	0.955	0.086	0.089
$\gamma$	0	0.0168	0.039	0.051	0.072
$\lambda$	0.2	0.066	0.196	0.086	0.007
$\mu$	0	-0.131	0.018	0.086	0.093
$\sigma$	1	0.921	1.002	0.103	0.126
$\gamma$	0.5	0.511	0.467	0.107	0.104
$\lambda$	0.2	0.383	0.192	0.128	0.069
$\mu$	0	-0.317	-0.024	0.068	0.264
$\sigma$	1	0.666	1.248	0.096	0.331
$\gamma$	1	1.047	0.772	0.138	0.109
$\lambda$	0.2	0.773	0.364	0.191	0.316
$\mu$	0	-0.399	-0.528	0.058	0.805
$\sigma$	1	0.388	1.780	0.087	0.877
$\gamma$	1.5	1.492	0.897	0.159	0.085
$\lambda$	0.2	1.317	0.916	0.253	0.937

Table 1: For each combination of  $\mu$ ,  $\sigma$  and  $\gamma$ , we generate 100 samples of  $\text{SGEV}(\mu, \sigma, \gamma, \lambda = 0.2)$  and calculate the PWM and ML estimations, together with the respective mean and standard deviation (sd).

$\lambda$  nearest to zero.

## 6 Conclusions

A new class of distributions, called SGEV distributions, has been proposed as a generalization of the classical GEV distribution, thanks to the introduction of a new parameter  $\lambda$ . Estimators of the parameters and quantiles of the SGEV distribution have been derived by ML and PWM methods. As mentioned in the above sections, the SGEV distribution is more flexible than the classical GEV distribution and so, can fit better maxima when dealing with small maxima samples or when the maxima come from small size



Parameter	True value	Mean for ML	Mean for PWM	sd for ML	sd for PWM
$\mu$	0	0.240	-0.021	0.278	0.100
$\sigma$	1	0.913	0.976	0.227	0.094
$\gamma$	-0.5	-0.389	-0.444	0.370	0.075
$\lambda$	0.5	0.029	0.503	0.424	0.028
$\mu$	0	0.209	0.003	0.107	0.104
$\sigma$	1	0.986	0.951	0.084	0.091
$\gamma$	0	0.012	0.046	0.089	0.075
$\lambda$	0.5	0.232	0.496	0.110	0.006
$\mu$	0	0.013	0.016	0.080	0.099
$\sigma$	1	0.983	0.995	0.089	0.134
$\gamma$	0.5	0.502	0.466	0.104	0.112
$\lambda$	0.5	0.491	0.493	0.125	0.050
$\mu$	0	-0.143	-0.126	0.112	0.292
$\sigma$	1	0.837	1.170	0.131	0.272
$\gamma$	1	0.997	0.760	0.133	0.113
$\lambda$	0.5	0.787	0.763	0.218	0.493
$\mu$	0	-0.212	-0.917	0.107	1.439
$\sigma$	1	0.654	1.937	0.170	1.019
$\gamma$	1.5	1.490	0.875	0.154	0.092
$\lambda$	0.5	1.033	1.819	0.313	1.968

Table 2: For each combination of  $\mu$ ,  $\sigma$  and  $\gamma$ , we generate 100 samples of  $\text{SGEV}(\mu, \sigma, \gamma, \lambda = 0.5)$  and calculate the PWM and ML estimations, together with the respective mean and standard deviation (sd).

Parameter	True value	Mean for ML	Mean for PWM	sd for ML	sd for PWM
$\mu$	0	0.022	0.006	0.445	0.094
$\sigma$	1	1.060	0.951	0.353	0.128
$\gamma$	-0.5	-0.349	-0.439	0.786	0.094
$\lambda$	1	0.124	1.005	1.330	0.027
$\mu$	0	0.241	0.009	0.249	0.089
$\sigma$	1	1.048	0.965	0.177	0.100
$\gamma$	0	0.196	0.031	0.424	0.069
$\lambda$	1	0.383	0.997	0.331	0.004
$\mu$	0	0.157	-0.014	0.107	0.105
$\sigma$	1	1.081	1.043	0.116	0.173
$\gamma$	0.5	0.513	0.454	0.128	0.117
$\lambda$	1	0.702	1.004	0.186	0.044
$\mu$	0	0.011	-0.273	0.074	0.429
$\sigma$	1	0.988	1.326	0.106	0.444
$\gamma$	1	1.028	0.763	0.132	0.119
$\lambda$	1	0.958	1.408	0.162	0.830
$\mu$	0	-0.100	-2.445	0.085	1.914
$\sigma$	1	0.827	2.334	0.166	1.825
$\gamma$	1.5	1.514	0.894	0.169	0.077
$\lambda$	1	1.262	3.123	0.219	2.416

Table 3: For each combination of  $\mu$ ,  $\sigma$  and  $\gamma$ , we generate 100 samples of  $\text{SGEV}(\mu, \sigma, \gamma, \lambda = 1)$  and calculate the PWM and ML estimations, together with the respective mean and standard deviation (sd).

Parameter	True value	Mean for ML	Mean for PWM	sd for ML	sd for PWM
$\mu$	0	0.137	-0.006	0.272	0.098
$\sigma$	1	0.936	0.935	0.222	0.119
$\gamma$	-0.5	-0.523	-0.421	1.005	0.085
$\lambda$	1.5	0.185	1.507	0.808	0.0198
$\mu$	0	0.179	0.009	0.345	0.100
$\sigma$	1	1.029	0.956	0.192	0.126
$\gamma$	0	0.591	0.027	0.746	0.069
$\lambda$	1.5	0.596	1.497	0.755	0.005
$\mu$	0	0.185	0.011	0.262	0.124
$\sigma$	1	1.114	1.021	0.180	0.206
$\gamma$	0.5	0.694	0.471	0.407	0.123
$\lambda$	1.5	1.089	1.508	1.991	0.058
$\mu$	0	0.138	-0.461	0.103	0.676
$\sigma$	1	1.104	1.389	0.154	0.522
$\gamma$	1	1.041	0.770	0.158	0.121
$\lambda$	1.5	1.222	2.219	0.188	1.473
$\mu$	0	0.006	-2.905	0.070	2.740
$\sigma$	1	0.969	2.732	0.134	1.403
$\gamma$	1.5	1.509	0.904	0.171	0.079
$\lambda$	1.5	1.501	3.127	0.193	2.492

Table 4: For each combination of  $\mu$ ,  $\sigma$  and  $\gamma$ , we compute 100 samples of  $\text{SGEV}(\mu, \sigma, \gamma, \lambda = 1.5)$  and calculate the PWM and ML estimations, together with the respective mean and standard deviation (sd).

Model	Method	$\mu$	$\sigma$	$\gamma$	$\lambda$
SGEV	ML	119.41	50.22	0.35	-0.15
	PWM	107.23	49.99	0.30	0.05
GEV	ML	111.31	47.39	0.35	-
	PWM	112.47	50.57	0.27	-

Table 5: Estimations of  $\mu$ ,  $\sigma$ ,  $\gamma$  and  $\lambda$  by ML and PWM methods for the Fort Collins data set, obtained by fitting a SGEV or a GEV distribution.

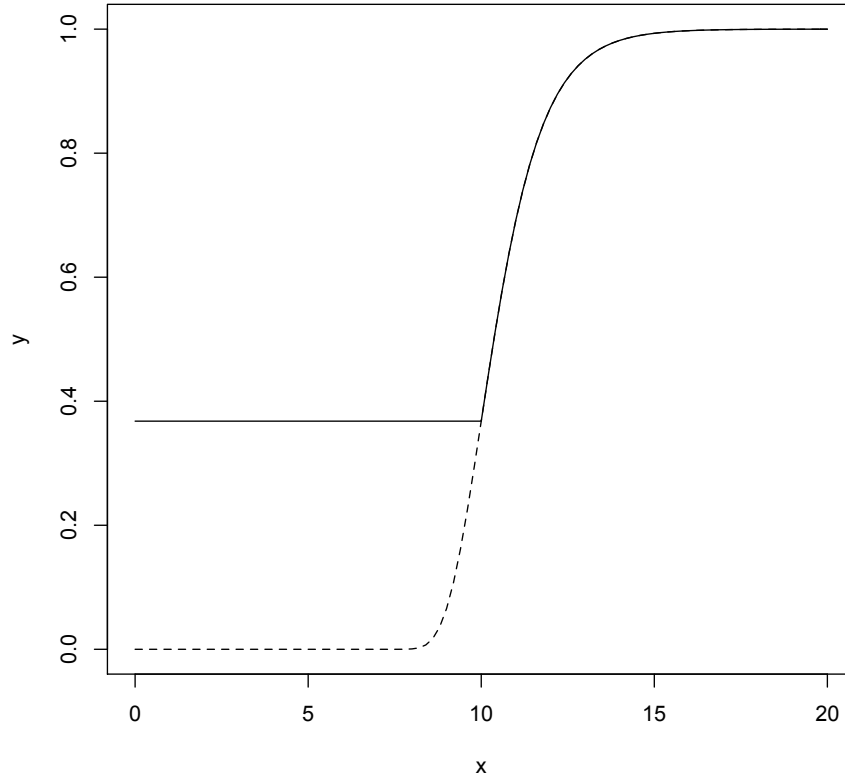


Figure 1: Cumulative distribution function of  $M$  if  $X_i \sim GPD(\mu = 10, \sigma = 1, \gamma = 0)$  and  $\lambda = 1$  (full line) of the Remark 1 and cumulative distribution function of the GEV distribution (dashed line).

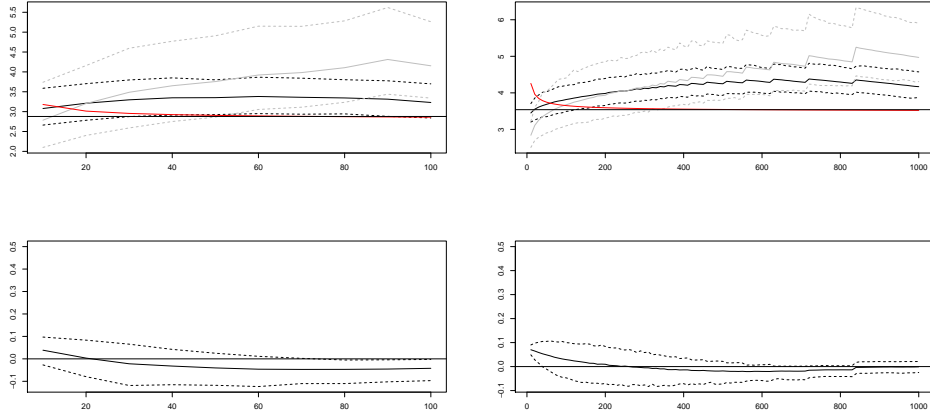


Figure 2: *Left hand*: Top row: mean, calculated on 500 samples of 500 realisations by a  $\mathcal{N}(0, 1)$  distribution, of the estimations of a  $1 - p$  quantile obtained by fitting a GEV (grey line) or a SGEV (dark line) with 95% empirical confidence bounds, as a function of the block size ( $x$  abscisse). The horizontal line represents the true quantile and the red line the quantile obtained by the block maxima method,  $z_p^{BM}$ . Bottom row: estimation of the skew parameter  $\lambda$ ;  $p = 1/500$ . *Right hand*: Same as left hand but based on 500 samples of 5000 realizations of a  $\mathcal{N}(0, 1)$  distribution and estimating a  $p$  quantile with  $p = 1/5000$ .

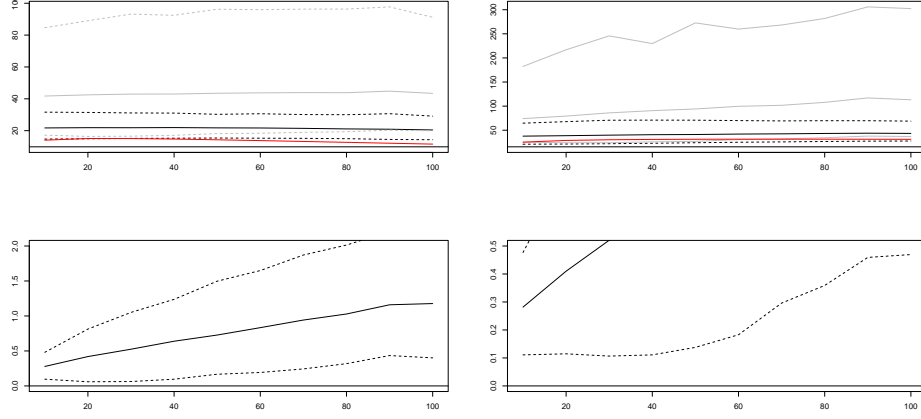


Figure 3: *Left hand:* The same as left hand in Figure 2 but with a Log-gamma(2,2) distribution ( $\gamma = 1/4$ ). *Right hand:* Same as left hand but with  $p = 1/2500$ .

blocks. Two interpretations of the SGEV distribution can be given. First, the cdf can be viewed as the cdf of a maximum over a random block size when the maximum is taken over exceedances (see Remark 1). Secondly, the SGEV distribution can be seen as the limiting distribution of the maxima when the original data are not independent. In that case, the parameter  $\lambda$  is just  $\theta - 1$ , where  $\theta$  is the extremal index. This means that the estimator of  $\lambda$ , obtained by the PWM method, can be looked at as a new and efficient estimator of the extremal index  $\theta$ .

Moreover, the SGEV could be used to find the right block size, the one such that  $\lambda$  is nearest to zero.

### Acknowledgements

This work has been partially supported by the MIRACCLE-GICC project, the French Research National Agency (ANR) under the reference ANR-08-BLAN-0314-01, the EU-FP7 “ACQWA” Project ([www.acqwa.ch](http://www.acqwa.ch)) under contract Nr 212250, the ANR-MOPERA project, the PEPER-GIS project, the ANR-McSim project and the research chair *Actuariat responsable* sponsored by Generali.

Authors would like to thank two anonymous referees for their helpful comments that led to an improvement of the present paper (we especially thank for Remarks 4 and 5).

## A Asymptotic covariance matrix

The asymptotic covariance matrix of the estimator vector obtained by the probability weighted moments method (see Theorem 1) can be written as

$$\Omega = \frac{(\lambda + 1)^2 \sigma^2}{\gamma^2} \begin{pmatrix} \Omega_{0,0} & \Omega_{0,1} & \Omega_{0,2} & \Omega_{0,3} \\ \Omega_{0,1} & \Omega_{1,1} & \Omega_{1,2} & \Omega_{1,3} \\ \Omega_{0,2} & \Omega_{1,2} & \Omega_{2,2} & \Omega_{2,3} \\ \Omega_{0,3} & \Omega_{1,3} & \Omega_{2,3} & \Omega_{3,3} \end{pmatrix}$$

and the elements on the main diagonal of  $\Omega$  can be specified as follows

$$\begin{aligned} \Omega_{0,0} &= (\lambda + 1)^{2\gamma} \left[ \Gamma(1 - 2\gamma) K \left( \frac{\lambda}{\lambda + 1} \right) - \Gamma^2(1 - \gamma) \right], \\ \Omega_{1,1} &= (2\lambda + 2)^{2\gamma} \left[ \Gamma(1 - 2\gamma) K \left( \frac{2\lambda + 1}{2\lambda + 2} \right) - \Gamma^2(1 - \gamma) \right], \\ \Omega_{2,2} &= (3\lambda + 3)^{2\gamma} \left[ \Gamma(1 - 2\gamma) K \left( \frac{3\lambda + 2}{3\lambda + 3} \right) - \Gamma^2(1 - \gamma) \right], \\ \Omega_{3,3} &= (4\lambda + 4)^{2\gamma} \left[ \Gamma(1 - 2\gamma) K \left( \frac{4\lambda + 3}{4\lambda + 4} \right) - \Gamma^2(1 - \gamma) \right], \end{aligned}$$

where  $K$  and  $\Gamma$  are respectively the hypergeometric and gamma functions. The remaining elements of the matrix are

$$\begin{aligned} \Omega_{a,a+1} &= \frac{1}{2} \left( ((a + 2)(\lambda + 1))^{2\gamma} \Gamma(1 - 2\gamma) K \left( \frac{\lambda(a + 1) + a}{(\lambda + 1)(a + 2)} \right) \right. \\ &\quad - (\lambda(a + 2) + a + 1)^\gamma \Gamma(1 - 2\gamma) K \left( \frac{(\lambda + 1)(a + 1)}{\lambda(a + 2) + a + 1} \right) \\ &\quad + 2((\lambda + 1)(a + 1))^\gamma \times [(\lambda(a + 2) + a + 1)^\gamma \\ &\quad \left. ((\lambda + 1)(a + 2))^\gamma \right] \Gamma^2(1 - \gamma) \Big), \end{aligned}$$

$$\begin{aligned}
\Omega_{a,a+2} = & \frac{1}{2} \left( ((a+3)(\lambda+1))^{2\gamma} \Gamma(1-2\gamma) K \left( \frac{\lambda(a+1)+a}{(\lambda+1)(a+3)} \right) \right. \\
& - (\lambda(a+3)+a+2)^\gamma \Gamma(1-2\gamma) K \left( \frac{(\lambda+1)(a+1)}{\lambda(a+3)+a+2} \right) \\
& + 2((\lambda+1)(a+1))^\gamma \times [(\lambda(a+3)+a+2)^\gamma \\
& \left. ((\lambda+1)(a+3))^\gamma] \Gamma^2(1-\gamma) \right),
\end{aligned}$$

$$\begin{aligned}
\Omega_{a,a+3} = & \frac{1}{2} \left( ((a+4)(\lambda+1))^{2\gamma} \Gamma(1-2\gamma) K \left( \frac{\lambda(a+1)+a}{(\lambda+1)(a+4)} \right) \right. \\
& - (\lambda(a+4)+a+3)^\gamma \Gamma(1-2\gamma) K \left( \frac{(\lambda+1)(a+1)}{\lambda(a+4)+a+3} \right) \\
& + 2((\lambda+1)(a+1))^\gamma \times [(\lambda(a+4)+a+3)^\gamma \\
& \left. ((\lambda+1)(a+4))^\gamma] \Gamma^2(1-\gamma) \right).
\end{aligned}$$

## References

- [1] Allard, D. and Naveau, P. (2007), A New Spatial Skew-Normal Random Field Model. *Communications in Statistics: Theory and Methodology*, **39**(9), 1821-1834.
- [2] Azzalini, A. (1985), A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171-178.
- [3] Azzalini, A. (1986), Further results on a class of distributions which includes the normal ones. *Statistica*, **46**, 199-208.
- [4] Azzalini, A. (2005), The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, **32**, 159-188.
- [5] Azzalini, A. (2013), *The Skew-Normal and Related Families*, Cambridge.
- [6] Beirlant J., Goegebeur Y., Segers J. and Teugels J. (2004), *Statistics of Extremes: Theory and Applications*, Wiley Series in Probability and Statistics.



- [7] Chang, S-M. and Genton, M. G. (2007), Extreme Value Distributions for the Skew-Symmetric Family of Distributions. *Communications in Statistics: Theory and Methodology*, **36**(9), 1705-1717.
- [8] Diebolt, J., Guillou, A., Naveau, P. and Ribereau, P. (2008), Improving Probability-Weighted Moments Methods for the Generalized Extreme Value Distribution. *REVSTAT*, **6**(1), 1-18.
- [9] Embrechts, P. Klüppelberg, C. and Mikosch, T. (2003), *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag.
- [10] Flecher, C., Naveau, P. and Allard, D. (2009a), Estimating the Closed Skew-Normal distributions parameters using weighted moments, *Statistics and Probability Letters*, **79**(19), 1977-1984.
- [11] Flecher, C., Naveau, P., Allard, D. and Brisson, N. (2009b), A Stochastic Daily Weather Generator for Skewed Data, *Water Resources Research*, in review.
- [12] Genton, M. G. (2004), *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Chapman & Hall.
- [13] Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J. R. (1979), Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, **15**(5), 1049-1054.
- [14] Guillou, A., Naveau, P., Diebolt, J. and Ribereau, P. (2009), Return level bounds for discrete and continuous random variables. *Test*, **18** (3), 584-604.
- [15] Gupta, A. K., Chang, F. C., and Huang, W. J. (2002), Some skewsymmetric models. *Random Operators and Stochastic Equations*, **10**, 133-140.
- [16] Hallin, M., Ley, C. (2012a), Skew-Symmetric Distributions and Fisher Information the double sin of the skew-normal. *arXiv:1209.4177*.
- [17] Hallin, M., Ley, C. (2012b) Skew-symmetric distributions and Fisher information - a tale of two densities. *Bernoulli*, **18** (3), 747-763.
- [18] Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985), Estimation of the Generalized Extreme-Value distribution by the method of probability-weighted moments. *Technometrics*, **27**, 251-261.

- [19] Jenkinson, A.F. (1955), The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorology Society*, **87**, 145-158.
- [20] Lysenko, N., Roy, P. and Waeber, R. (2009), Multivariate extremes of generalized skew-normal distributions. *Statistics and Probability Letters*, **79**, 525-533.
- [21] Nadarajah, S. and Kotz, S. (2006). Skew distributions generated from different families. *Acta Appl. Math.* **91**(1), 1-37.
- [22] Nadarajah, S. and Kotz, S. (2007a), Skew Models I, *Acta Applicandae Mathematicae*, **98**(1), 1-28.
- [23] Nadarajah, S. and Kotz, S. (2007b), Skew Models II, *Acta Applicandae Mathematicae*, **98**(1), 29-46.
- [24] Naveau, P., Genton, M. G. and Shen, X. (2005), A skewed Kalman filter, *Journal of multivariate Analysis*, **94**, 382-400.
- [25] Pickands, J. (1975), Statistical inference using extreme order statistics, *Annals of Statistics*, **3**, 119-131.
- [26] Sartori, N. (2006), Bias prevention of maximum likelihood estimates for scalar skew normal and skew  $t$  distributions. *Journal of Statistical Planning and Inference*, **136**, 4259-4275.
- [27] Smith, R. (1985), Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, **72**, 67-90.
- [28] Vernic, R. (2006), Multivariate skew-normal distributions with applications in insurance. *Insurance: Mathematics and economics*, **38**, 413-426.